

WHAT IS CLAIMED IS:

1. A speech recognition apparatus for recognizing speech by comparing composite acoustic models adapted to noise and speaker
5 with a feature vector series extracted from an uttered speech, comprising:

a storing section for previously storing each representative acoustic model selected as a representative of acoustic models belonging to one of groups, each of said groups being formed beforehand
10 by classifying a large number of acoustic models on a basis of a similarity, difference models of each group obtained from difference between said acoustic models belonging to one of said groups and said representative acoustic model of said identical group, and group information for corresponding said representative acoustic
15 models with said difference models every said identical group,

a generating section for generating each noise adaptive representative acoustic model of said each group by noise-adaptation executed to said representative acoustic model of said each group stored in said storing section;

20 a generating section for generating each composite acoustic model of said each group by composition of said difference model and said noise adaptive representative acoustic model using said group information;

a renewal model generating section for generating noise and
25 speaker adaptive acoustic models by performing a speaker-adaptation of said composite acoustic model every identical group, using the feature vector series obtained from the uttered speech; and

a model renewal section for replacing said difference models of said each group by renewal difference models which are generated by taking differences between said noise and speaker adaptive acoustic models and said noise adaptive representative acoustic models selected via said group information;

wherein a speech recognition is performed by comparing the feature vector series extracted from the uttered speech to be recognized with said composite acoustic model adapted to noise and speaker, and

wherein said composite acoustic model adapted to noise and speaker is generated by composition of said renewal difference model and said noise adaptive representative acoustic model, which is generated by a noise-adaptation of said representative acoustic model of said group including said renewal difference model selected via said group information.

2. A speech recognition apparatus for recognizing speech by comparing composite acoustic models adapted to noise and speaker with a feature vector series extracted from an uttered speech, comprising:

a storing section for previously storing each representative acoustic model selected as a representative of acoustic models belonging to one of groups, each of said groups being formed beforehand by classifying a large number of acoustic models on a basis of a similarity, difference models of each group obtained from difference between said acoustic models belonging to one of said groups and said representative acoustic model of said identical group, and

group information for corresponding said representative acoustic models with said difference models every said identical group,

a generating section for generating each noise adaptive representative acoustic model of said each group by noise-adaptation executed to said representative acoustic model of said each group stored in said storing section;

a generating section for generating each composite acoustic model of said each group by composition of said difference model and said noise adaptive representative acoustic model using said group information;

a recognition processing section for recognizing speech by comparing said composite acoustic models generated in said generating section for composite acoustic models with the feature vector series extracted from the uttered speech to be recognized;

a renewal model generating section for generating noise and speaker adaptive acoustic models by performing a speaker-adaptation of said composite acoustic model every identical group, using the feature vector series obtained from the uttered speech; and

a model renewal section for replacing said difference models of said each group by renewal difference models which are generated by taking differences between said noise and speaker adaptive acoustic models and said noise adaptive representative acoustic models selected via said group information;

wherein said recognition processing section performs a speech recognition by comparing the feature vector series extracted from the uttered speech to be recognized with said composite acoustic model adapted to noise and speaker generated by composition of said

noise adaptive representative acoustic model generated by
noise-adaptation of said representative acoustic model of each group
including said renewal difference model selected with said group
information and said renewal difference model renewed by said renewal
5 model generating section and said model renewal section, every
repetition of the speech recognition.

3. The speech recognition apparatus according to claim 1 or 2,
wherein

10 said model renewal section repeats to change the group
including said noise and speaker adaptive acoustic model of the
group information based on a similarity of said noise and speaker
adaptive acoustic model and said noise adaptive representative
acoustic model, every generation of said renewal difference model,
15 and

said difference model stored in said storing section is renewed
with the difference between said noise and speaker adaptive acoustic
model and said noise adaptive representative acoustic model of the
group including said noise and speaker adaptive acoustic model
20 selected based on said renewed group information.

4. A speech recognition method for recognizing speech by comparing
a set of composite acoustic models adapted to noise and speaker
with a feature vector series extracted from an uttered speech,
25 comprising the steps of:

previously storing, in a storing section, each representative
acoustic model selected as a representative of acoustic models

belonging to one of groups, each of said groups being formed beforehand by classing a large number of acoustic models on a basis of a similarity, difference models of each group obtained from difference between said acoustic models belonging to one of said groups and said
5 representative acoustic model of said identical group, and group information for corresponding said representative acoustic models with said difference models every said identical group;

generating each noise adaptive representative acoustic model of said each group by noise-adaptation executed to said
10 representative acoustic model of said each group stored in the storing section;

generating each composite acoustic model of said each group by composition of said difference model and said noise adaptive representative acoustic model using said group information;

15 generating noise and speaker adaptive acoustic models by performing a speaker-adaptation of said composite acoustic model every identical group, using the feature vector series obtained from the uttered speech; and

replacing said stored difference models of said each group
20 by renewal difference models which are generated by taking differences between said noise and speaker adaptive acoustic models and said noise adaptive representative acoustic models selected via said group information;

wherein said speech recognition is performed by comparing
25 the feature vector series extracted from the uttered speech to be recognized with said composite acoustic model adapted to noise and speaker, and

wherein said composite acoustic model adapted to noise and speaker is generated by composition of said renewal difference model and said noise adaptive representative acoustic model, which is generated by a noise-adaptation of said representative acoustic model of said group including said renewal difference model selected via said group information.

5. A speech recognition method for recognizing speech by comparing a set of composite acoustic models adapted to noise and speaker with a feature vector series extracted from an uttered speech, comprising the steps of:

previously storing, in a storing section, each representative acoustic model selected as a representative of acoustic models belonging to one of groups, each of said groups being formed beforehand by classing a large number of acoustic models on a basis of a similarity, difference models of each group obtained from difference between said acoustic models belonging to one of said groups and said representative acoustic model of said identical group, and group information for corresponding said representative acoustic models with said difference models every said identical group,

generating each noise adaptive representative acoustic model of said each group by noise-adaptation executed to said representative acoustic model of said each group stored in the storing section;

generating each composite acoustic model of said each group by composition of said difference model and said noise adaptive representative acoustic model using said group information;

recognizing a speech by comparing said composite acoustic models generated in said generating step for composite acoustic models with the feature vector series extracted from the uttered speech to be recognized;

5 generating noise and speaker adaptive acoustic models by performing a speaker-adaptation of said composite acoustic model every identical group, using the feature vector series obtained from the uttered speech; and

 replacing said stored difference models of said each group
10 by renewal difference models which are generated by taking differences between said noise and speaker adaptive acoustic models and said noise adaptive representative acoustic models selected via said group information;

 wherein said recognition processing step performs a speech
15 recognition by comparing the feature vector series extracted from the uttered speech to be recognized with said composite acoustic model adapted to noise and speaker generated by composition of said noise adaptive representative acoustic model generated by
noise-adaptation of said representative acoustic model of each group
20 including said renewal difference model selected with said group information and said renewal difference model renewed by said noise and speaker adaptive acoustic models generating step and said difference models replacing step, every repetition of the speech recognition.

25

6. The speech recognition method according to claim 4 or 5, wherein said difference models replacing step repeats to change the

group including said noise and speaker adaptive acoustic model of the group information based on a similarity of said noise and speaker adaptive acoustic model and said noise adaptive representative acoustic model, every generation of said renewal difference model,

5 and

said difference model stored in said storing section is renewed with the difference between said noise and speaker adaptive acoustic model and said noise adaptive representative acoustic model of the group including said noise and speaker adaptive acoustic model

10 selected based on said renewed group information.